



# Project Status Report : SAMGrid



- ♦ **SAMGrid Requirements from Stakeholders**
- ♦ **SAMGrid Deployments**
- ♦ **SAMGrid Development**
- ♦ **SAMGrid Matchup & Questions**



# SAMGrid Requirements from Stakeholders



## ♦ CDF

- Use of SAMGrid on CAF; 25% of analysis offsite (due now)
- Use of SAMGrid for farm production – new in plan
- Quota-enabled permanent disk storage mechanism
- Tools for concatenation of output files
- Quota-enabled tape storage of user-generated analysis files
- VO policy implementations across remote centers

## ♦ CMS – potential stakeholder

- Support for 'service challenges' as well as 'data challenges'
- $10^7$  evts/mo sim/reco/analyzed – due now
- Analysis on the Grid: Nov 04
- Joint projects desired: common SE's based on SRM; architecture effort w/in PPDG
- Interoperability: OSG & LCG



# SAMGrid Requirements from Stakeholders



## ♦ DØ

- **JIM for MC (now) & RECO (Sept) & analysis (Mar 05)**
- **Full linux operation (Mar 05)**
- **2<sup>nd</sup> gen monitoring**
- **LCG interoperability**
- **Tape migration procedures (for next drive gen)**
- **SRM interface (timetable from CDF)**
- **Security/VO man as needed by team**
- **Program to reduce operational load**



# SAMGrid Requirements from Stakeholders



## ♦ MINOS

- Need to use a URL with no file movement and then reorder list before opening – currently, exp't requirement to read files in order (this may change). → use SAM/dCache station instead of SAM cache station for now
- Would like to explore AFS disk as SAM cache
- Would like to add field DetectorType in schema
- No GRID requirements as yet
- ORACLE machine (Linux!) expected to be deployed by Aug1 – drives user deployment (tentative goal end 04)
- Other storage issues:
  - competition for dCache resources in stken
  - want to continue offsite dcap/dccp access w/o auth.
  - would migrate to 200GB tapes if procedure available



# SAMGrid Deployments



## ♦ CDF

- **SAM offsite:**  
Datasets offsite at CNAF & Acad Sinica, in progress to SDSC  
More working DCAFs with SAM DH: Rutgers, Japan, Korea, ...  
Level of usage disappointing: onsite sufficient, less data  
Production import of MC: Toronto this week, then move to more sites
- **SAM on CAF:**  
Full suite of desired project recovery tools not available  
In use by small number of ordinary users ( at least 1!)
- **CDF H Stream reprocessing** – done (possibly up to small # of files), lessons learned being produced; farm prototype proposed by end July
- **JIM deployment issues:**  
Slow because stretched thin; can live with human scheduling
- **New schema went to production this week: fairly smooth**



# SAMGrid Deployments



## ♦ CMS

- No requirement for deployment from CMS yet
- Ongoing discussions with respect to SAM schema & DH prototype could possibly lead to some form of deployments in future

## ♦ DØ

- JIM for MC (& RECO)  
Lyon, Manchester, Wisconsin – in production  
Langston, Lancaster – done  
GridKA, OSCER – in progress (GridKA req VDT migration)  
Prague, Imperial, RAL, FNAL farm – to start next  
UTA, WestGrid, some smaller stations – needed by Sept
- JIM for analysis – not started yet
- New schema in production 6/1/04; new dbs tests in progress



# SAMGrid Deployments



## ♦ MINOS

- 2 stations: one w/ SAM cache, one w/ dCache
- Offline framework runs SAM projects using C++ API
- Jobs run on fnalu from both stations
- Sam data\_browsing tools on www\_numi (still minor fixes from schema mismatch being worked on)



# SAMGrid Development

## [Slide from Feb 04]



- ♦ **Active Subprojects:** C++ API, DBServer, JIM, H Stream Reco for CDF, Caching, Chains&Links, CDF DFC, Test Harness, Linux deploy of DBServers, Config Man
- ♦ **Planned Subprojects:** Request system, Autodest, Further monitoring (MIS)
- ♦ **Related Subprojects:** d0tools, SBIR II, Condor mods, workflow packages for CDF & D0, Authorization & Accounting
- ♦ **Recently completed Subprojects:** Python API, V5.1 Schema Design, Batch Adapter, D0 Online dcache TDP, 1<sup>st</sup> Gen Monitoring Tools, Data Dimensions Grammar





# SAMGrid Development [Updated]



- ♦ **Active Subprojects:** C++ API, DBServer, JIM,, Caching, Chains&Links, CDF DFC, SAM on CAF, Further monitoring (MIS), Request system, Autodest
- ♦ **Planned Subprojects:** Followups for test harness, DBServer, Hstream RECO
- ♦ **Related Subprojects:** d0tools, SBIR II, Condor mods, workflow packages for CDF & D0, Authorization & Accounting
- ♦ **Recently completed Subprojects:** H Stream Reco for CDF, Linux deploy of DBServers, Config Man, Test Harness, Python API, V5.1 Schema Design, Batch Adapter, D0 Online dcache TDP, 1<sup>st</sup> Gen Monitoring Tools, Data Dimensions Grammar

Subproject	Status	Exp't Req
<b>C++ API</b>	<b>SAM_HOST_NAME issue in progress</b>	<b>CDF deploy</b>
<b>DBServer</b>	<b>est 6/30; followups autodev &amp; request sys</b>	<b>CDF deploy D0 deploy</b>
<b>JIM</b>	<b>VDT mig ~ done; ongoing tests &amp; cleanups</b>	<b>CDF deploy D0 deploy</b>
<b>Hstream RECO</b>	<b>Done</b>	<b>CDF deploy</b>
<b>Caching/SRM</b>	<b>Tech plan in progress</b>	<b>CDF(disk store), D0</b>
<b>Chains &amp; Links</b>	<b>Stopped</b>	<b>CDF, D0 (new dim)</b>
<b>CDF DFC</b>	<b>In progress</b>	<b>CDF deploy</b>
<b>MIS</b>	<b>In progress</b>	<b>D0 (shifters, ops)</b>
<b>Request systems</b>	<b>Just starting</b>	<b>Farm &amp; MC ops</b>
<b>Autodev</b>	<b>Just starting</b>	<b>CDF deploy</b>



# SAMGrid Requirements from Stakeholders: Matchups



## ♦ CDF

- Use of SAMGrid on CAF; 25% of analysis offsite (due now)  
**In progress:**  
<http://projects.fnal.gov/samgrid/cdf/cdfdeploy.html>
- Use of SAMGrid for farm production – new in plan  
**In progress: email exchange**
- Quota-enabled permanent disk storage mechanism  
**In progress: cache/SRM subproject, Fella interim design**
- Tools for concatenation of output files  
**In progress: Hstream RECO**
- Quota-enabled tape storage of user-generated analysis files  
**Storage in progress; quota-enabling not started**
- VO policy implementations across remote centers  
**Not started**



# SAMGrid Requirements from Stakeholders: Matchups



## ♦ CMS

- Support for 'service challenges' as well as 'data challenges'
- $10^7$  evts/mo sim/reco/analyzed – due now
- Analysis on the Grid: Nov 04
- Joint projects desired: common SE's based on SRM; architecture effort w/in PPDG
- Interoperability: OSG & LCG

Not definite enough to evaluate yet. Recent activities: Lyon talk @ ARDA session; Veseli/Lyon session w/ Guglielmo et al.



# SAMGrid Requirements from Stakeholders: Matchups



## ♦ DØ

- JIM for MC (now) & RECO (Sept) & analysis (Mar 05)  
**In progress**
- Full linux operation (Mar 05)  
**Linux testing going well, no showstopper for this**
- 2<sup>nd</sup> gen monitoring  
**In progress: MIS project**
- LCG interoperability  
**Demo of client functionality on LCG @ Apr workshop**
- Tape migration procedures (for next drive gen)  
**Not started**
- SRM interface (timetable from CDF)  
**In progress: cache/SRM subproject**
- Security/VO man as needed by team  
**In progress: VO privilege participation**
- Program to reduce operational load  
**In progress: MIS project, doc week, general directives**
- Resource brokering  
**Not started**



# SAMGrid Requirements from Stakeholders: Matchups



## ♦ MINOS

- Need to use a URL with no file movement and then reorder list before opening – currently, exp't requirement to read files in order (this may change). → use SAM/dCache station instead of SAM cache station for now  
**Current functionality suffices for now**
- Would like to explore AFS disk as SAM cache  
**Not started**
- Would like to add field DetectorType in schema  
**Not started**
- No GRID requirements as yet
- ORACLE machine (Linux!) expected to be deployed by Aug1 – drives user deployment  
**SAMGrid effort will be required in August time frame**